Security Analysis and Improvement of a Multi-Factor Biometric-Based Remote Authentication Scheme

Sirapat Boonkrong

***Abstract*—Since the advent of the Internet technology, many remote user authentication protocols have been designed and developed which makes secure communication between a user and a server possible over an insecure channel. One of the first multi-factor biometric-based authentication scheme is known as the Li-Hwang protocol. Since then, there have been many analyses and improvements on the protocol to make it more secure. The most recently enhanced multi-factor biometric-based authentication has been proposed by Park *et al.* However, this research shows that several weaknesses can still be found including an attack on the integrity of a protocol message, the possibility of a replay attack and the lack of proof of message authenticity. This research, therefore, proposes an improved multi-factor biometric-based authentication scheme which has also been proved secure, using the GNY logic, against the mentioned attacks.**

***Index Terms*—authentication, biometrics, cryptanalysis, cryptography, multi-factor authentication, protocol.**

I. Introduction

THE Internet technology has allowed for convenience in many services and applications, most notably in communication between users and servers. Unfortunately, such ease of communication leads to the transmission of data over insecure channels. This causes many security challenges, including confidentiality and integrity of messages. In the security context, adversaries are usually considered and assumed to have control over communication channels between users and servers. In order to reduce the risk, a user and a server must prove their identity to one another. They also need to establish a session key for a more secure communication session. This is where authentication protocols are required.

Authentication is a method that is a part of access control mechanism. It is used for proving or confirming the identity of an entity. Remote user authentication is, therefore, an approach used to verify the identity of a user before accessing a service on a network or the Internet. One of the most common method used for remote user authentication is one-factor authentication, using password only. However, it has been accounted that the password-only schemes have been a target of many forms of attacks [1], including brute force attack, password dictionary attack, key logger attack and eavesdropping or shoulder surfing. In order to overcome this issue, several biometric-based user authentication protocols have been proposed [2], [3], [4], [5], [6]. In other words, instead of using a password to prove the identity of a user, biometric information such as fingerprints and irises is used. This is because they are considered to be a unique identifier for each person and difficult to forge. Thus, biometric-based authentication has been claimed to be more secure than the password-only scheme in the sense that it is less easily stolen and provides a stronger defense against repudiation [7].

Even though biometrics seem to provide better security than passwords, it is not a completely vulnerability-free solution. Some biometrics have been successfully forged in the past. Therefore, it is recommended that a second factor of authentication is still needed. In other words, many protocols have adopted a mechanism known as multi-factor authentication [8], [9], [10] instead of using a biometric on its own. Multi-factor authentication is a method where users need to provide more than one legitimate credential (or factor of authentication) for the purposes of identity verification and confirmation. Those authentication factors may include a password, biometric information and a smart card.

One of the first biometric-based remote user authentication scheme using smart cards protocols was proposed by Li and Hwang in 2010 [2]. The authors had applied a one-way hash function, biometric verification and a smart card to their protocol and claimed that the protocol was secure. However, a year later, Das [3] illustrated that there were some design flaws in the Li-Hwang scheme. As a result, the author proposed an enhanced protocol by incorporating more components, including a password and a random nonce into the protocol. In 2012, several weaknesses were found in the Das protocol by [4]. Those vulnerabilities included user impersonation attack, server masquerading attack, password guessing attack as well as the lack of mutual authentication between a user and a server. Therefore, some improvements to the protocol were proposed in order to increase the security of the protocol [4].

Although several improvements had been made to the original Li-Hwang's multi-factor authentication protocol, Cao and Ge were able to point out that the An protocol still contained some weaknesses, including replay attacks and server masquerading [5]. Cao and Ge, therefore, designed an improved multi-factor biometric-based authentication protocol to address those issues. Most recently, in 2017, Park *et al.* [6] demonstrated that it was still possible for adversaries to carry out attacks on the Cao-Ge protocol, including identity guessing attack and server masquerading attack.

It can be seen that since the Li-Hwang protocol in 2010, there have been many attempts made in designing and developing a more secure multi-factor biometric-based remote authentication protocol. Unfortunately, several vulnerabilities in the scheme proposed by Park *et al.* have been found by

Manuscript received April 4, 2019; revised August 18, 2019.

S. Boonkrong is with the School of Information Technology and with DIGITECH, Suranaree University of Technology, Nakhon Ratchasima, 30000 Thailand email: sirapat@sut.ac.th.

this research. This paper, therefore, discusses the weaknesses in the Park *et al.*'s scheme and proposes an improved multi-factor biometric-based remote authentication protocol with enhanced security. An analysis on security and correctness, which applies the Gong-Needham-Yahalom or the GNY logic [11], is provided together with the analysis of the efficiency.

The rest of the paper is organised as follows. Section 2 provides the description of the protocol proposed by Park *et al.* and an overview of the GNY logic. The vulnerabilities of the Park *et al.*'s protocol are explained in Section 3. Section 4 provides the design and security analysis of an improved multi-factor biometric-based remote authentication protocol. The efficiency analysis of the proposed protocol is given in Section 5. Section 6 then concludes the paper.

II. Background Knowledge

Preliminary knowledge that is related to the work done in this paper is explained in this section. This is comprised of the description of the Park *et al.*'s remote user authentication protocol and an introduction to the logic of GNY.

A. *Park* et al.*'s remote user authentication protocol*

Park *et al.* proposed a multi-factor biometric-based remote user authentication protocol whose aim was for a user and a server to be mutually authenticated with a shared session key established at the end of the protocol run. The protocol consists of three main phases. They are the registration phase, the login phase and the authentication phase. The notations of the scheme [6] are described in Table I.

TABLE I
Notations in the Park *et al.*'s protocol

Notation	Meaning
C_i	User i
R_i	Trusted Registration Centre i
S_i	Remote Server i
ID_i	Actual Identity of C_i
VID_i	Virtual Identity of C_i
DID_i	Dynamic Identity of C_i
PW_i	Password of C_i
B_i	Biometric Template of C_i
SC_i	Smart Card of User C_i
A_i	Adversary i
X_S	Secret Key of S_i
x	Master Key of R_i
K	Random number for Registration of C_i
R_C	Random number generated by C_i
R_S	Random number generated by S_i
$\|\|$	Concatenate Operation
$\oplus$	Bitwise XOR Operation
$h()$	Secure Hash Function
$H()$	Bio-Hashing Function
y_i	A Unique Number of C_i generated by R_i
T_i	Timestamp

Registration phase

In this phase, a user C_i must complete his or her registration at the registration centre R_i before accessing any service on the server S_i. C_i begins the process by choosing his or her identity ID_i and a password PW_i. He or she then imprints biometric information B_i and carries out the following steps.

(R1) A user C_i chooses a random number K, and computes $(PW_i \oplus K)$ and $(H(B_i) \oplus K)$. C_i then submits $ID_i, (PW_i \oplus K), (H(B_i) \oplus K)$ to R_i via a secure channel.

(R2) R_i chooses a unique number y_i of C_i and computes: $f_i = h(H(B_i) \oplus K)$, $r_i = h(PW_i \oplus K) \oplus f_i$, $e_i = h(ID_i||X_S) \oplus r_i$, $VID_i = h(y_i||X_S) \oplus ID_i \oplus h(PW_i \oplus K||h(H(B_i) \oplus K))$, $Z_i = y_i \oplus h(x)$ and $G_i = h(h(ID_i||XS))$.

(R3) R_i creates an entry of this identity ID_i and the virtual identity VID_i for this record.

(R4) R_i stores $\{VID_i, h(), H(), f_i, e_i, Z_i, G_i\}$ on a smart card SC_i. The smart card is then delivered to C_i via a secure channel.

(R5) Once C_i receives the smart card SC_i, he or she stores the random number K on the smart card.

Login phase

Once registered, the user C_i can attempt to log into a remote server S_i. This is the phase where such process takes place. When logging in, C_i carries out the following steps using his or her smart card SC_i.

(L1) C_i imprints his or her biometric information B_i and computes $H(B_i)$. The smart card SC_i then computes $h(H(B_i) \oplus K)$ and compares it with f_i already on the card. If it is valid, the next steps are carried out.

(L2) C_i chooses a random number R_C and inputs (ID_i, PW_i, R_C) into the smart card, SC_i, which then computes:
$r'_i = h(PW_i \oplus K) \oplus f_i$, $M_1 = e_i \oplus r'_i$, $M_2 = M_1 \oplus R_C$, $M_3 = h(M_1||R_C||T_1)$.
C_i generates a dynamic identity DID_i by computing $DID_i = VID_i \oplus h(h(y_i||X_S)||T_1)$, where $h(y_i||X_S) = VID_i \oplus ID_i \oplus h(PW_i \oplus K||h(H(B_i) \oplus K))$.

(L3) The user C_i send the login request to S_i with the message: $\{DID_i, Z_i, M_2, M_3, T_1\}$.

Authentication phase

In the authentication phase, the remote server S_i and user C_i attempt to prove their identity to one another. The following steps take place during this authentication phase.

(A1) Having received the message $\{DID_i, Z_i, M_2, M_3, T_1\}$ from C_i, the server S_i verifies whether $T_1 - T \leq \Delta T$. If the verification fails, S_i stops the session. If the verification succeeds, S_i checks the validity of DID_i by comparing VID'_i with VID_i stored in the account database, where $VID'_i = DID_i \oplus h(h(y_i||X_S)||T_1)$ and $y_i = Z_i \oplus h(x)$.

(A2) If the checking fails, S_i stops the session. Otherwise, S_i computes $M_4 = h(ID_i||X_S)$, $M_5 = M_2 \oplus M_4$, and checks whether $M_3 = h(M_4||M_5||T_1)$ or not.

(A3) If the verification fails, S_i stops the session. Otherwise, S_i randomly chooses a number R_S and computes $M_6 = M_4 \oplus R_S$, $M_7 = h(M_4||R_S||T_2)$. S_i then sends the message $\{M_6, M_7, T_2\}$ to C_i.

(A4) C_i verifies $T_2 - T \leq \Delta T$. If the verification fails, C_i terminates the sessions. If the verification succeeds, C_i computes $M_8 = M_1 \oplus M_6$. C_i then checks the validity of $M_7 = h(M_1||M_8||T_2)$.

(A5) If the checking fails, C_i terminates the session. Otherwise, C_i computes $M_9 = h(M_1||R_C||M_8||T_2)$ and sends the message $\{M_9, T_3\}$ to S_i. C_i then computes a session key $SK = h(R_C||M_8||T_2||T_3||M_9)$.

(A6) S_i verifies the timestamp $T_3 - T \leq \Delta T$. If the verification fails, S_i terminates the session. If it succeeds, S_i validates whether $M_9 = h(M_4||M_5||R_S||T_2)$. If not, S_i terminates the session, else S_i computes the session key $SK' = h(M_5||R_S||T_2||T_3||M_9)$. S_i sends $h(SK')$ to C_i.

(A7) C_i compares the received $h(SK')$ with his or her own $h(SK)$. If they match, C_i and S_i share the same session key SK.

B. The Logic of GNY

The logic of Gong-Needham-Yahalom or the logic of GNY was introduced in [11] as an improved method on the BAN logic [12] which has been used in the analyses of many existing multi-factor biometric remote authentication schemes, including the analysis of the Park *et al.*'s protocol. Since the GNY logic is an improved version of the BAN logic which contains several limitations [11], the GNY logic will be used to analyse the improved authentication protocol proposed in this paper.

In general, the logic of GNY provides a way to analyse and investigate the components of protocol messages when they are transmitted from the source to destination. It allows for the analysis of cryptographic protocols step-by-step according to its specified postulates. The six categories of postulates of the GNY logic are the T or being-told postulates, the P or possession postulates, the F or freshness postulates, the R or recognisability postulates, the I or Message Interpretation postulates and the J or Jurisdiction postulates.

Suppose that P and Q are protocol principals. According to the GNY logic, the main notations are presented in Table II. For more detail of the GNY notations and postulates, see [11].

TABLE II
MAIN NOTATIONS OF THE GNY LOGIC

Notation	Meaning
$P \triangleleft X$	P is told formula X.
$P \ni X$	P possesses formula X.
$P \mid\sim X$	P once conveyed formula X.
$P \mid\equiv \sharp X$	P believes that formula X is fresh.
$P \mid\equiv \phi X$	P believes that formula X is recognisable.
$P \mid\equiv P \overset{S}{\leftrightarrow} Q$	P believes that S is a suitable secret for P and Q. Only P and Q know S.
$P \triangleleft *X$	P is told formula X that he did not convey previously.
$P \mid\equiv C$	If C is a statement, P believes that the statement C holds.

III. CRYPTANALYSIS OF PARK *et al.*'S AUTHENTICATION PROTOCOL

In this section, the security of Park *et al.*'s multi-factor biometric-based remote user authentication protocol is analysed. As stated earlier, Park *et al.* had analysed and found weaknesses in the Cao-Ge authentication scheme [6]. They, therefore, designed an improved protocol which was claimed to have overcome the problems. Unfortunately, we have found that the Park *et al.*'s protocol still contains security vulnerabilities of its own.

When carrying out the analysis of the protocol, it is assumed that an adversary can have the following capabilites. These capabilities are consistent with those stated in [6] when the authors analysed the Cao-Ge authentication protocol.

The first capability is that an adversary A_i has the total control over the communication channel during all three phases of the protocol - registration, login and authentication. In other words, A_i can intercept, insert, modify and delete any message that is transmitted via the communication channel. The second capability is that A_i can either steal the user's smart card or obtain the user's password. The third capability is that the adversary A_i can extract the information stored on the smart card.

A. Cryptanalysis of the registration phase

Park *et al.* claimed that their registration step was an improvement on the Cao and Ge's registration phase. However, a couple of vulnerabilities still exist as follows. First of all, the first message of the Park *et al.*'s registration process is sent by C_i to R_i as:

$$C_i \rightarrow R_i : ID_i, (PW_i \oplus K), (H(B_i) \oplus K).$$

A vulnerability that can be found here is that if an adversary A_i intercepts and makes a modification to the message, the registration centre R_i will not be able to detect it. In other words, A_i can make changes to any of the components of the message such as changing ID_i to ID_j. R_i will not be able to detect those changes at all. The reason for not being able to detect the modification is that the message does not contain any mechanism, specifically a cryptographic hash function, which can assist in such detection. If, for example, ID_i is changed to ID_j without being realised by the recipient, the consequence is that R_i will store wrong information on the smart card as well as in its own database. Therefore, the vulnerability of the first message of the registration phase is the lack of message integrity detection mechanism.

The second message of the Park *et al.*'s registration process is sent by R_i to C_i via a smart card as:

$$R_i \rightarrow C_i : VID_i, h(), H(), f_i, e_i, Z_i, G_i.$$

The first problem that can be seen here is that out of all the seven components on the smart card SC_i, only changes on two components f_i and r_i can be detected by C_i if occurred. This is because, by the description of the Park *et al.*'s protocol, these two components are computed with the items known and generated by C_i. This means that the rest of the components give a similar weakness to the first message. That is, by the assumption stated above, A_i can get hold of SC_i and extract information from it. The adversary A_i can also make changes to the information on SC_i. Again, there is no mechanism that C_i can use to check the integrity of the data at all. This is, therefore, one vulnerability of this registration message.

The second weakness of this message is the fact that a replay attack is possible. The reason is that there is no evidence to suggest that the components sent to C_i are newly generated by R_i every time the message is constructed. More simply, the recipient C_i has no way to prove whether any of the received components are fresh. This means that an adversary A_i can again extract the information on the smart card and re-use it for a replay attack. The recipient will not be able to know whether the received information is a replay or newly generated.

Furthermore, based on the description of the registration process by Park *et al.*, there is a lack of a component, information or a process to convince C_i that the information on the smart card has really been conveyed by the registration centre R_i, which could lead to a man-in-the-middle attack. This is the third vulnerability of this message of the registration process.

B. Cryptanalysis of the login and authentication phase

The aims of the login and authentication phase are for the user C_i to log into the remote server S_i, while mutually authenticating one another, and for them to establish a new session key. Although Park *et al.* claimed to have mitigated the vulnerabilities of the Cao and Ge's login and authentication phase, several weaknesses can still be found. They can be explained as follows. The login and authentication phase begins with the message:

$$C_i \rightarrow S_i : DID_i, Z_i, M_2, M_3, T_1.$$

The first weakness found in this message is that if an adversary makes changes to the message, the remote server S_i will not be able to detect them. This means if any of the components of the message is modified, S_i will not be able to find out because there is no mechanism, such as a cryptographic hash function, that can assist in the detection process. If, for example, DID_i is changed without being realised by S_i, the consequence is that the user C_i will not be able to log into the server. Therefore, the first weakness of the message is the lack of message integrity detection mechanism.

The second message of the login and authentication phase is sent by the remote server S_i to the user C_i as:

$$S_i \rightarrow C_i : M_6, M_7, T_2.$$

Similar to the above message, there is a lack of message integrity detection mechanism, which makes it possible for an attacker to make changes to the message without being detected. Another problem with this message is that there is no component that can be recognised by the user C_i. In other words, this message contains no components that has been generated and sent by C_i, which S_i uses, as a part of this message, to produce this reply. This is a lack of challenge-and-response mechanism. There is no proof to C_i that this message is really from S_i, the other entity that C_i is communicating with. The implication is that a man-in-the-middle attack is possible due to this vulnerability.

Another problem that has been discovered in this stage of the protocol is the lack of mutual understanding of session key possession. In other words, the steps (A5) - (A7) in the Park *et al*'s login and authentication are the session key establishment process. Within these three steps, only the server S_i proves to the user C_i that it is holding the same session key. However, there is no process that allows C_i to prove to S_i that he or she is also in possession of the same session key. This means that at the end of (A7), C_i believes that S_i is holding the same session key, but S_i does not have the knowledge that C_i is holding the same session key.

Even though the remote server S_i proves to the user C_i that it is in possession of the same session key SK, the way that Park *et al.* designed to accomplish this task is for S_i to send $h(SK)$ to C_i. We believe that this method can cause a security problem. Park *et al.* did not mention which hash function $h()$ to use in this case. Even that, if MD5 is used, there is a possibility of a rainbow table attack [13] as well as a collision attack [14]. The former can ultimately lead to an attacker being able to discover SK.

IV. Proposed Remote Authentication Protocol

We propose a multi-factor biometric-based remote authentication protocol that we believe can overcome the the security problems explained in the previous section. In the proposed protocol, we apply a cryptographic hash function for the message integrity detection purposes. It is also ensured that a replay attack is not feasible by adding a fresh component to protocol messages. Moreover, a challenge-and-response mechanism is used so that man-in-the-middle attack is not possible. Finally, an extra message is added to the authentication protocol so that mutual understanding of the session key can be confirmed. Similar to the Park *et al.*'s protocol and the Cao-Ge protocol, the proposed protocol consists of two main parts. They are the registration phase and the login and authentication phase, and can be explained as follows.

A. Registration Phase

The proposed registration phase still contains the same structure as the Park *et al.*'s protocol with a couple of improvements that help overcome the previously mentioned weaknesses. A message integrity detection mechanism or a cryptographic hash function $h()$ has been added to both registration messages. In both messages of this phase, a new component called N_{C_i} and N_{R_i} or the nonce value generated by the user C_i amd registration centre R_i have been included to ensure the freshness of the message. This can prevent a replay attack. Moreover, these new components are used as a challenge in the challenge-and-response mechanism which has been lacking in the Park *et al.*'s protocol. The proposed registration phase can be written as follows.

$$\begin{array}{ll} \text{(M1)}\, C_i \rightarrow R_i : & ID_i, N_{C_i}, (PW_i \oplus K), (H(B_i) \oplus K), \\ & h(ID_i, (PW_i \oplus K), (H(B_i) \oplus K)) \\ \text{(M2)}\, R_i \rightarrow C_i : & VID_i, N_{C_i}, N_{R_i}, h(), H(), f_i, e_i, Z_i, G_i, \\ & h(VID_i, N_{C_i}, N_{R_i}, h(), H(), f_i, e_i, Z_i, G_i) \end{array}$$

It can be seen in the messages M1 and M2 that the components N_{C_i} and N_{R_i} have now been incorporated in the messages. The main purpose is to ensure the freshness of the messages since N_{C_i} and N_{R_i} are newly generated each time a new registration is carried out. Therefore, a replay attack is no longer possible. Furthermore, the component N_{C_i} is included in the message M2 as part of the response from R_i to C_i. This is so that C_i knows that the message is really from the entity that he or she intends to communicate with. This can be thought of as a challenge-and-response mechanism. In addition, a cryptographic hash function $h()$ is now applied on both messages. C_i and R_i will be able to check the integrity of the received message accordingly.

B. Login and Authentication Phase

The proposed login and authentication phase contains several new components and mechanisms that are necessary for mitigating the vulnerabilities of the Park *et al.*'s scheme. Firstly, new components N_{C_i} and N_{S_i} or the nonce values generated by the user and remote server are introduced to the first message and fourth message of this phase, respectively. They are untilised to guarantee the freshness of the message and is a part of the challenge-and-response mechanism, which is lacking in the existing scheme. Secondly, a message

integrity mechanism or a cryptographic hash function $h()$ is added to all the messages. Thirdly, an extra component and a mechanism are added to the login and authentication phase in order to ensure that both principals C_i and S_i know and believe that the other party holds the same session key SK. The proposed login and authentication phase is as follows.

(M1) $C_i \rightarrow S_i$: $DID_i, N_{C_i}, Z_i, M_2, M_3, T_1, h(DID_i, N_{C_i}, Z_i, M_2, M_3, T_1)$
(M2) $S_i \rightarrow C_i$: $N_{C_i}+1, M_6, M_7, T_2, h(N_{C_i}+1, M_6, M_7, T_2)$
(M3) $C_i \rightarrow S_i$ $\{N_{C_i}+2\}_{SK}, M_9, T_3, h(\{N_{C_i}+2\}_{SK}, M_9, T_3)$
(M4) $S_i \rightarrow C_i$: $\{N_{C_i}+3, N_{S_i}\}_{SK'}, T_4, h(\{N_{C_i}+3, N_{S_i}\}_{SK'}, T_4)$

It can be seen in the proposed protocol that the main functionality of the login and authentication protocol of the Park *et al.*'s scheme is still there. What has been done is to improve its security in three folds. The first is that the recipient of a message, whether it is the user C_i or the remote server S_i will be able to check the integrity of the message, because of the introduction of a cryptographic hash function $h()$.

Secondly, in messages M1 to M4, the nonce N_{C_i} is generated and used in the following way. The user C_i generates the nonce N_{C_i} and sends it along with other components to the server S_i, comprising the first message M1. Having received and processed M1, S_i increments the nonce N_{C_i} by 1 and sends it back to the user, together with other components, forming the message M2. The purpose of this process is so that the user C_i knows and believes that the principal that responds to him or her is really the principal he or she intends to send the first message to, namely the remote server S_i. Similarly, in the messages M3 and M4, the nonce N_{C_i} is incremented by 1 in order to ensure that the recipient of the message believes that the message is from the intended and expected entity, namely C_i and S_i, respectively.

Thirdly, messages M3 and M4 have been designed with the purpose of the mutual understanding of a new session key. In other words, in M3, after C_i has generated a session key SK, he or she uses it to encrypt the nonce $N_{C_i}+2$ and sends it to the server S_i. At this stage, S_i will also produce the same session key. If the key is generated correctly, S_i will be able to decrypt the ciphertext $\{N_{C_i}+2\}_{SK}$. The server S_i verifies whether the result of the decryption process is something expected or something that has been computed from the nonce $N_{C_i}+1$ sent by S_i previously. If it is the expected result, S_i knows and believes that C_i is possessing the same session key. S_i will also be able to encrypt $N_{C_i}+3$ and a newly generated nonce N_{S_i} before transmitting them to C_i in the message M4.

Upon receiving M4, C_i decrypts the ciphertext of $N_{C_i}+3$ and N_{S_i} and verifies whether $N_{C_i}+3$ is the expected value. If so, C_i knows and believes that S_i is now holding the same session key. This completes the process of the mutual understanding of session key. This mechanism is lacking in the Park *et al.*'s protocol.

V. Analysis

This section contains the analysis of both the registration and authentication phases of the proposed multi-factor biometric-based remote authentication protocol. The capabilities of adversaries are assumed to be the same as those from the cryptanalysis of the Park *et al.*'s scheme, stated in Section III-A. First, the security and correctness of the proposed protocol are proved and analysed using the logic of GNY. Second, the general analysis of security on the proposed scheme as well as the comparison of security features with the Cao-Ge and Park *et al.*'s schemes are carried out. Third, the performance of the proposed scheme is compared to the existing protocols.

A. GNY analysis on the proposed protocol

In this section, the proposed protocol is analysed using the logic of GNY, which is a formal method used for proving the security and correctness of cryptographic protocols, as explained in Section II-B. It should be noted that the analysis method applied here is different from the BAN logic used in the Park *et al.*'s scheme. This is because several flaws have been found in the BAN logic [11] and the logic of GNY is its refinement.

Before the analysis of the proposed protocol is carried out, it should be pointed out that its security goals are as follows. Note that these security goals are kept exactly the same as the goals of the Park *et al.*'s protocol, for consistency.

(G1) $C_i \mid\equiv C_i \overset{SK}{\leftrightarrow} S_i$ — C_i believes that the key SK is shared between C_i and S_i.
(G2) $S_i \mid\equiv S_i \overset{SK}{\leftrightarrow} C_i$ — S_i believes that the key SK is shared between S_i and C_i.
(G3) $C_i \mid\equiv S_i \mid\equiv S_i \overset{SK}{\leftrightarrow} C_i$ — C_i believes S_i believes that the key SK is shared between S_i and C_i.
(G4) $S_i \mid\equiv C_i \mid\equiv S_i \overset{SK}{\leftrightarrow} S_i$ — S_i believes C_i believes that the key SK is shared between C_i and S_i.

Furthermore, Park *et al.* made the following assumptions before their analysis was carried out. No additional assumptions are made here in our analysis of the proposed protocol. This is so that the analysis is consistent with how it was done by Park *et al.*

$C_i \mid\equiv C_i \overset{h(y_i||X_S)}{\longleftrightarrow} S_i$
$C_i \mid\equiv C_i \overset{h(ID_i||X_S)}{\longleftrightarrow} S_i$
$S_i \mid\equiv \sharp R_C$
$S_i \ni x$

$S_i \mid\equiv C_i \overset{h(y_i||X_S)}{\longleftrightarrow} S_i$
$S_i \mid\equiv C_i \overset{h(ID_i||X_S)}{\longleftrightarrow} S_i$
$C_i \mid\equiv \sharp R_S$
$S_i \ni X_S$

We first analyse the registration phase, then the login and authentication phase is analysed.

1) Analysis of the registration phase: First of all, the two messages of the registration phase are written in the GNY idealised form as follows.

(M1) $R_i \triangleleft$ $*ID_i, *N_{C_i}, *(PW_i \oplus K), *(H(B_i) \oplus K), *h(*ID_i, *(PW_i \oplus K), *(H(B_i) \oplus K))$
(M2) $C_i \triangleleft$ $*VID_i, *N_{C_i}, *N_{R_i}, *h(), *H(), *f_i, *e_i, *Z_i, *G_i, *h(*VID_i, *N_{C_i}, *N_{R_i}, *h(), *H(), *f_i, *e_i, *Z_i, *G_i)$

The GNY analysis on the first message M1 can now be carried out.

Message M1:

Applying T1: We get

$$R_i \triangleleft ID_i, N_{C_i}, (PW_i \oplus K), (H(B_i) \oplus K), h(ID_i, (PW_i \oplus K), (H(B_i) \oplus K))$$

This means that the registration centre R_i hears the message M1 and all of its components.

Applying P1: We get

$$R_i \ni ID_i, N_{C_i}, (PW_i \oplus K), (H(B_i) \oplus K), h(ID_i, (PW_i \oplus K), (H(B_i) \oplus K))$$

This means that R_i possesses the message M1 and all of its components.

Applying P4: P4 states that if R_i possesses the components of the message then R_i is able to compute the hash value of the message. This implies that R_i can carry out message integrity checking on the received message.

Applying F1: We obtain

$$R_i \mid\equiv \sharp(ID_i, N_{C_i}, (PW_i \oplus K), (H(B_i) \oplus K), h(ID_i, (PW_i \oplus K), (H(B_i) \oplus K)))$$

Here, the registration centre R_i checks its database. If the component ID_i is not found then R_i believes that ID_i is fresh. Therefore, R_i believes that the whole message M1 is fresh, i.e., the message is not a replay attack.

At the end of the analysis on the first registration message, it is learned that the registration centre R_i now possesses the components sent by the user C_i. R_i has checked the integrity of the message and believes that the received message is not a replay.

The GNY analysis is now carried out on the second message M2 in the registration phase.

Message M1:

Applying T1: We get

$$C_i \triangleleft VID_i, N_{C_i}, N_{R_i}, h(), H(), f_i, e_i, Z_i, G_i, h(VID_i, N_{C_i}, N_{R_i}, h(), H(), f_i, e_i, Z_i, G_i)$$

This means that the user C_i hears the message M2 and all of its components sent to him or her by the registration centre R_i.

Applying P1: We get

$$C_i \ni VID_i, N_{C_i}, N_{R_i}, h(), H(), f_i, e_i, Z_i, G_i, h(VID_i, N_{C_i}, N_{R_i}, h(), H(), f_i, e_i, Z_i, G_i)$$

This means that C_i possesses the message M2 and all of its components.

Applying P4: P4 states that if C_i possesses the components of the message then C_i is able to compute the hash value of the message. This implies that C_i can carry out message integrity checking on the received message.

Applying F1: We obtain

$$C_i \mid\equiv \sharp(VID_i, N_{C_i}, N_{R_i}, h(), H(), f_i, e_i, Z_i, G_i, h(VID_i, N_{C_i}, N_{R_i}, h(), H(), f_i, e_i, Z_i, G_i))$$

Due to the component N_{R_i} that has been added to the protocol, the user C_i knows that N_{R_i} is a newly generated component. Therefore, C_i believes that the message M2 must also be fresh, i.e., the message does not constitute a replay attack.

Applying R1: Here C_i believes that the received component f_i is recognisable because C_i possesses $h()$ and $H(B_i) \oplus K$, and $f_i = h(H(B_i) \oplus K)$. Moreover, C_i also recognises the nonce N_{C_i} that has been added to the registration phase, because it is the component generated and transmitted by C_i in the message M1. In other words, a challenge-and-response mechanism is being used here. Therefore, by applying the GNY postulate R1, we obtain

$$C_i \mid\equiv \phi(VID_i, N_{C_i}, N_{R_i}, h(), H(), f_i, e_i, Z_i, G_i, h(VID_i, N_{C_i}, N_{R_i}, h(), H(), f_i, e_i, Z_i, G_i))$$

This means that the message M2 is considered recognisable by the user C_i.

Applying I3: This is a message interpretation postulate of the logic of GNY. Here, we apply the fact that the user C_i has received and possessed the message M2 and its components. C_i believes that the biometric information $H((B_i) \oplus K)$ is shared between himself or herself and the registration centre R_i. C_i also believes that the message is fresh. Therefore, we can now obtain

$$C_i \mid\equiv R_i \mid\sim VID_i, N_{R_i}, h(), H(), f_i, e_i, Z_i, G_i, C_i \mid\equiv R_i \mid\sim h(VID_i, N_{R_i}, h(), H(), f_i, e_i, Z_i, G_i)$$

By applying the I3 postulate, it can now be interpreted that the user C_i believes that the registration centre R_i is really the one who has constructed and conveyed the message M2.

Applying J1: This is to apply a jurisdiction postulate of the GNY logic. Carrying on from the I3 postulate above, the user C_i believes that R_i is the one who has conveyed the message. Hence, C_i believes that R_i also believes and has some jurisdiction over the message M2 and all the components. Therefore, the application of the J1 postulate gives us

$$C_i \mid\equiv VID_i, N_{R_i}, h(), H(), f_i, e_i, Z_i, G_i, C_i \mid\equiv h(VID_i, N_{R_i}, h(), H(), f_i, e_i, Z_i, G_i)$$

This implies that the user C_i believes in the message M2 and all of its components as well.

The analysis and proof of correctness of the messages in the registration phase ends here. What has been learned from the GNY analysis on both messages is two folds. Firstly, the registration centre R_i possesses all the components sent by the user C_i and has been able to check the integrity of the received message, too. Secondly, the user C_i is sure that the message M2 from R_i is freshly generated. More importantly, he or she believes that the message and its components have really been transmitted by the registration centre R_i. These are the results of the proposed mechanisms in challenge-and-response, cryptographic hash function and nonce values.

Next, the login and authentication phase is analysed and proved by the logic of GNY.

2) Analysis of the login and authentication phase: As proposed earlier, the login and authentication phase for the multi-factor biometric-based remote authentication protocol contains four messages, which can be idealised in the format of GNY as follows.

(M1) $S_i \triangleleft \ *DID_i, *N_{C_i}, *Z_i, *M_2, *M_3, *T_1, *h(*DID_i, *N_{C_i}, *Z_i, *M_2, *M_3, *T_1)$

(M2) $C_i \triangleleft \ *N_{C_i}+1, *M_6, *M_7, *T_2, *h(*N_{C_i}+1, *M_6, *M_7, *T_2)$

(M3) $S_i \triangleleft \ *\{*N_{C_i}+2\}_{SK}, *M_9, *T_3, *h(*\{*N_{C_i}+2\}_{SK}, *M_9, *T_3)$

(M4) $C_i \triangleleft \ *\{*N_{C_i}+3, *N_{S_i}\}_{SK'}, *T_4, *h(*\{*N_{C_i}+3, *N_{S_i}\}_{SK'}, *T_4)$

The analysis and proof of correctness using the logic of GNY is now carried out on the login and authentication phase. The analysis begins with the first message M1.

Message M1:

Applying T1: We get

$$S_i \triangleleft DID_i, N_{C_i}, Z_i, M_2, *M_3, T_1, h(DID_i, N_{C_i}, Z_i, M_2, M_3, T_1)$$

This means that the remote server S_i hears the message M1 and all of its components.

Applying P1: We get

$$S_i \ni DID_i, N_{C_i}, Z_i, M_2, *M_3, T_1, h(DID_i, N_{C_i}, Z_i, M_2, M_3, T_1)$$

This means that S_i possesses the message M1 and all of its components.

Applying P4: P4 states that if S_i possesses the components of the message then S_i is able to compute the hash value of the message. This implies that S_i can carry out message integrity checking on the received message.

Applying F1: Here, the remote server S_i can check whether the timestamp T1 is new or fresh. If S_i believes that the timestamp T1 is fresh, then we obtain

$$S_i \mid\equiv \sharp(DID_i, N_{C_i}, Z_i, M_2, *M_3, T_1, h(DID_i, N_{C_i}, Z_i, M_2, M_3, T_1))$$

This means that the remote server believes that the message M1 is fresh, i.e., the message does not constitute a replay attack.

Applying R1: At this stage, the remote server S_i computes VID_i' from $h(H(y_i||X_S)||T_1) \oplus DID_i$. If it matches with VID_i stored in its account database, then S_i can believe that VID_i is recognisable. Hence, DID_i is also recognisable by S_i. That is, with the application of the recognisability postulate R1, we obtain

$$S_i \mid\equiv \phi(DID_i, N_{C_i}, Z_i, M_2, *M_3, T_1, h(DID_i, N_{C_i}, Z_i, M_2, M_3, T_1))$$

This means that the message M1 is considered recognisable by the remote server S_i.

At the end of the GNY analysis on the first message, it is learned that the remote server S_i has received and possessed the login message from the user C_i. The server can check the integrity of the message by using the additional mechanism in cryptographic hash function. Moreover, the server is certain that the received message is not a replay due to the use of timestamp as well as a proposed nonce value.

The second login and authentication message M2 is now analysed.

Message M2:

Applying T1: We get

$$C_i \triangleleft N_{C_i} + 1, M_6, M_7, T_2, h(N_{C_i} + 1, M_6, M_7, T_2)$$

This means that the user C_i hears the message M2 and all of its components.

Applying P1: We get

$$C_i \ni N_{C_i} + 1, M_6, M_7, T_2, h(N_{C_i} + 1, M_6, M_7, T_2)$$

This means that C_i possesses the message M2 and all of its components.

Applying P4: P4 states that if C_i possesses the components of the message then C_i is able to compute the hash value of the message. This implies that C_i can carry out message integrity checking on the received message.

Applying F1: Here, there are two components that C_i can use to check for freshness. The first is the timestamp component T_2. The second is the component R_S via the computation of the component M_6 and M_7. If C_i believes that both components have been freshly generated, then we obtain

$$C_i \mid\equiv \sharp(N_{C_i} + 1, M_6, M_7, T_2, h(N_{C_i} + 1, M_6, M_7, T_2))$$

This means that the user believes that the message M2 is fresh, i.e., the message is not a result of a replay attack.

Applying R1: At this stage, the user C_i can compute $N_{C_i} + 1$ in order to check whether this received component has been computed from his or her previously generated nonce N_{C_i}. If this is the case, it can be implied that C_i recognises the component $N_{C_i} + 1$. Therefore, we obtain

$$C_i \mid\equiv \phi(N_{C_i} + 1, M_6, M_7, T_2, h(N_{C_i} + 1, M_6, M_7, T_2))$$

This means that the message M2 is considered recognisable by the user C_i. In turn, it can be understood by C_i that the message M2 has been replied by the expected entity S_i. This is possible due to the proposed challenge-and-response mechanism.

Applying I3: A message interpretation postulate of the GNY logic is applied. Here, we make use of the fact that the user C_i has received and possessed the message M2 as well as all of its components. C_i also believes that the component M_6 is shared between himself or herself and the remote server S_i due to the component X_S used to computeM_4, which in turn is used to compute M_6. In addition, by the analysis of the postulate F1, C_i believes that the message M2 is fresh. Therefore, we can now obtain

$$C_i \mid\equiv S_i \mid\sim N_{C_i} + 1, M_6, M_7, T_2, h(N_{C_i} + 1, M_6, M_7, T_2)$$

By applying the I3 postulate, it can now be interpreted that the user C_i believes that the remote server S_i is really the one who has constructed and conveyed the message M2 to him or her.

Applying J1: A jurisdiction postulate of the GNY logic is next to be applied in the analysis. At this stage of the anlaysis, C_i believes that S_i is the one who has conveyed the message M2. Hence, C_i believes that S_i also believes and has some jurisdiction over the message M2 and all its components. Therefore, the application of the J1 postulate gives us

$$C_i \mid\equiv N_{C_i} + 1, M_6, M_7, T_2, h(N_{C_i} + 1, M_6, M_7, T_2)$$

This implies that the user C_i believes in the message M2 and all of its components as well.

The analysis of the second message of the login and authentication phase provides us with the fact that the user C_i has received and possessed the message and its components. C_i can also check the integrity of the message to ensure that nothing has been modified before the message is received by

him or her. Moreover, the proposed challenge-and-response mechanism helps C_i to believe that the message M2 has really be transmitted by the expected entity in S_i.

Next, the message M3 of the login and authentication phase is analysed. Note that by the description of the proposed protocol, the message M3 is transmitted by the user C_i to the remote server S_i after the new session key SK has been computed by the user.

Message M3:

Applying T1: We get

$$S_i \triangleleft \{N_{C_i} + 2\}_{SK}, M_9, T_3, \\ h(\{N_{C_i} + 2\}_{SK}, M_9, T_3)$$

This means that the remote server S_i hears the message M3 and all of its components.

Applying P1: We get

$$S_i \ni \{N_{C_i} + 2\}_{SK}, M_9, T_3, \\ h(\{N_{C_i} + 2\}_{SK}, M_9, T_3)$$

This means that S_i possesses the message M3 and all of its components.

Applying P4: P4 states that if S_i possesses the components of the message then S_i is able to compute the hash value of the message. This implies that S_i can carry out message integrity checking on the received message.

Applying F1: The remote server S_i checks whether or not the timestamp T_3 is fresh. If S_i believes that the component T_3 or the timestamp is fresh, then we obtain

$$S_i \mid\equiv \sharp(\{N_{C_i} + 2\}_{SK}, M_9, T_3, \\ h(\{N_{C_i} + 2\}_{SK}, M_9, T_3))$$

This means that the remote server S_i believes that the message M3 is fresh, i.e., the message is not a replay attack.

At this stage, according to the description of the proposed protocol, the remote server S_i computes its own session key SK', which is then used to decrypt the component $\{N_{C_i} + 2\}_{SK}$. Suppose that S_i has computed the session key, a postulate T3 can be applied.

Applying T3: After a session key SK' has been computed by the server S_i, we have the fact that $S_i \ni SK'$ or S_i possesses the session key SK'. Suppose that the session key SK' can be used to decrypt the message, then by applying the postulate T3 we obtain

$$S_i \triangleleft N_{C_i} + 2$$

Here, if the postulate P1 is applied again, we will obtain that $S_i \ni N_{C_i} + 2$, which means that S_i now possesses the component $N_{C_i} + 2$. The recognisability postulate of GNY can be applied next.

Applying R1: Once the component $\{N_{C_i} + 2\}_{SK}$ has been decrypted, S_i obtains $N_{C_i} + 2$. The server then checks whether this component has been computed from $\{N_{C_i} + 1\}$ which it has previously computed and transmitted to C_i in the message M2. If this is the case, the application of R1 gives us

$$S_i \mid\equiv \phi(\{N_{C_i} + 2\}_{SK}, M_9, T_3, \\ h(\{N_{C_i} + 2\}_{SK}, M_9, T_3))$$

This means that the message M3 is considered recognisable by the remote server S_i. In turn, it can be understood by S_i that the message M3 has been replied by the expected entity C_i. This is possible due to the proposed challenge-and-response mechanism.

Applying I3: From the fact that the server S_i has received and possessed the message M3 and its component, with the fact that S_i also believes that the received message is not a replay as well as the fact that the server believes that at least one component such as N_{C_i} is shared between itself and the user C_i, we can obtain

$$S_i \mid\equiv C_i \mid\sim \{N_{C_i} + 2\}_{SK}, M_9, T_3, \\ h(\{N_{C_i} + 2\}_{SK}, M_9, T_3)$$

By the application of the postulate I3, it can be interpreted that the remote server S_i believes that the user C_i is really the one who has constructed and transmitted the message M3.

Applying J1: We now apply a jurisdiction postulate of the GNY logic. At this stage of the anlaysis, S_i believes that C_i is the one who has conveyed the message M3. Hence, S_i believes that C_i also believes and has some jurisdiction over the message M3 and all its components. Therefore, the application of the J1 postulate gives us

$$S_i \mid\equiv \{N_{C_i} + 2\}_{SK}, M_9, T_3, \\ h(\{N_{C_i} + 2\}_{SK}, M_9, T_3)$$

This implies that the remote server S_i believes in the message M3 and all of its components as well.

After the proof of correctness and the GNY analysis on the message M3, it is learned that the server S_i has received and possessed the message and its components and has been able to ensure that the integrity of the message is intact by the use of the proposed mechanism, namely a cryptographic hash function. S_i also believes that this message is not a replay attack due to the use of a proposed nonce value. Furthermore, for the sake of the analysis, S_i has computed its own session key and has been able to decrypt the component $N_{C_i} + 2$. This implies that S_i can be certain that the message M3 is really from the expected entity in the user C_i. This has been made possible by the proposed challenge-and-response mechanism. In fact, by constructing and transmitting the message M3, C_i has shown to S_i that he or she has computed and possessed the same session key. S_i believes that this is the case because it has seen and recognised the component N_{C_i}, more specifically $N_{C_i} + 2$.

The message M4 of the login and authentication protocol is to be analysed by the logic of GNY next. Note that the purpose of the message M4 is for the remote server S_i to prove to the user C_i that it also possesses the same session key as the user. The analysis on the message is now carried out.

Message M4:

Applying T1: We get

$$C_i \triangleleft \{N_{C_i} + 3, N_{S_i}\}_{SK'}, T_4, \\ h(\{N_{C_i} + 3, N_{S_i}\}_{SK'}, T_4)$$

This means that the user C_i hears the message M4 and all of its components.

Applying P1: We get

$$C_i \ni \{N_{C_i} + 3, N_{S_i}\}_{SK'}, T_4, \\ h(\{N_{C_i} + 3, N_{S_i}\}_{SK'}, T_4)$$

This means that C_i possesses the message M4 and all of its components.

Applying P4: P4 states that if C_i possesses the components of the message then C_i is able to compute the hash value of the message. This implies that C_i can carry out message integrity checking on the received message.

Applying F1: The user C_i checks whether or not the received timestamp component T_4 is fresh. If C_i believes that the timestamp T_4 is fresh, then we obtain

$$C_i \mid\equiv \sharp(\{N_{C_i}+3, N_{S_i}\}_{SK'}, T_4, h(\{N_{C_i}+3, N_{S_i}\}_{SK'}, T_4))$$

This means that the user C_i believes that the message M4 is fresh, i.e., the message does not constitute a replay attack.

At this stage of the proposed authentication protocol, the user C_i attempts to decrypt the component $\{N_{C_i}+3, N_{S_i}\}_{SK'}$ using his or her previously computed session key SK. That is, the postulate T3 is applied in the analysis as follows.

Applying T3: As stated earlier, the user C_i possesses the session key SK, i.e., $C_i \ni SK$. Therefore, by applying T3, we obtain

$$C_i \lhd N_{C_i}+3, N_{S_i}$$

Here, if the postulate P1 is applied again, we will obtain that $C_i \ni N_{C_i}+3, N_{S_i}$, which means that C_i now possesses the components $N_{C_i}+3$ and N_{S_i}. The recognisability postulate of GNY can be applied next.

Applying R1: Once the nonce $\{N_{C_i}+3\}_{SK}$ has been decrypted, the user C_i obtains $N_{C_i}+3$. The user then examines whether or not this component has been computed from the nonce $\{N_{C_i}+2\}$ which he or she has previously computed and transmitted to the remote server S_i in the message M3. If this is true then the application of the GNY postulate R1 gives us

$$C_i \mid\equiv \phi(\{N_{C_i}+3, N_{S_i}\}_{SK'}, T_4, h(\{N_{C_i}+3, N_{S_i}\}_{SK'}, T_4))$$

This means that the message M4 is considered recognisable by the user C_i. In turn, it can be understood by C_i that the message M4 has been replied by the expected entity S_i. This is possible due to the proposed challenge-and-response mechanism. More importantly, the user C_i now believes that he or she and the remote server S_i now share the same session key SK. This is because S_i is able to encrypt the nonce $N_{C_i}+3$ and C_i is able to decrypt it correctly, too.

Applying I3: So far it has been learned that the user C_i has received and possessed the message M4 and all its components. C_i also believes that the received message is not a replay. Moreover, he or she believes that at least one component of the message, including the nonce N_{C_i} and more importantly the session key SK, is recognisable and shared between himself or herself and the server S_i. Therefore, by applying the message interpretation postulate I3, we can obtain

$$C_i \mid\equiv S_i \mid\sim \{N_{C_i}+3, N_{S_i}\}_{SK'}, T_4, h(\{N_{C_i}+3, N_{S_i}\}_{SK'}, T_4)$$

By applying I3, it can be interpreted that the user C_i believes that the remote server S_i or the expected entity is really the one who has conveyed the message M4. This also confirms that C_i believes that S_i possesses the same session key as him or her.

Applying J1: A jurisdiction postulate of the GNY logic is now applied to the message in the next step of the analysis. At this stage, C_i believes that S_i is the one who has conveyed the message M4. Hence, C_i believes that ${}_Si$ also believes and has some jurisdiction over the message M4 and all its components. Therefore, the application of J1 gives us

$$C_i \mid\equiv \{N_{C_i}+3, N_{S_i}\}_{SK'}, T_4, h(\{N_{C_i}+3, N_{S_i}\}_{SK'}, T_4)$$

This implies that the user C_i believes in the message M4 and all its components.

After the analysis and proof of correctness on the message M4, it can be seen that the user C_i has received and possessed the message and its components. The user has also been able to ensure that the message integrity can be checked by using the proposed mechanism in a cryptographic hash function. C_i also believes that this message is not a replay attack due to the use of a proposed nonce value. Moreover, C_i has been able to decrypt the component $N_{C_i}+3$, which implies that the message has really been conveyed by the expected principal, i.e., the remote server S_i. In addition, by constructing and transmitting the message M4, the server S_i has shown to the user C_i that it is possessing the same session key.

This ends the analysis on our proposed multi-factor biometric-based remote authentication protocol. It can be seen that the proposed protocol has achieved and accomplished the aims of overcoming the vulnerabilities of the Park *et al.*'s scheme. The proposed protocol has an added mechanism to help check the integrity of the message. The protocol uses additional nonce values to ensure the freshness of the messages. Hence, replay attacks have been mitigated. A chellenge-and-response mechanism has been applied so that each principal is able to identify that the other principal is really the expected entity. Lastly, both the user C_i and the remote server S_i can now believe that they really share the same session key at the end of the protocol run.

B. Security analysis

The cryptanalysis of the Park *et al.*'s authentication protocol was carried out in Section III, and an improved scheme was designed and logically analysed in Sections IV and V-A, respectively. This section, therefore, explains whether or not the vulnerabilities of the Park *et al.*'s protocol have been mitigated. The issues with the existing scheme include message falsification, replay attack, man-in-the-middle attack and lack of mutual understanding of session key.

1) Message falsification: The Park *et al.*'s authentication protocol has a problem of not being able to detect modifications to protocol messages. In the proposed protocol, it has become more difficult for an adversary to make changes to the messages without being detected. This is because all the messages in the registration, login and authentication phases have been designed to contain and apply a cryptographic hash function, $h()$. The hash function acts as a message integrity checking mechanism. That means if any changes were made to any of the protocol messages, they would be detected. Undetected message falsification is, therefore, not possible.

2) Replay attack: A replay is a possible attack on the Park *et al.*'s protocol in both the registration and login phases. The proposed protocol, therefore, introduces the use of nonce values in N_{Ci}, N_{Ri} and N_{Si} generated by the user, the registration centre and the remote server, respectively. These components are newly generated each time a message is created in order to ensure freshness of messages. Thus, if an old or replayed message were transmitted, it would be detected by the recipient.

3) Man-in-the-middle attack: In order to carry out a man-in-the-middle attack on the Park *et al.*'s authentication protocol, an adversary could intercept the protocol messages and forward them to the recipient, who would think that he or she was really communicating with a legitimate entity. However, in the proposed protocol, a challenge-and-response mechanism has been introduced in both registration and authentication phases. The challenge is simply the nonce values generated by the sending entity. The recipient can then reply by using the nonce as a part of his or her response. This mechanism can be seen in all the messages of the proposed protocol.

4) Lack of key mutual understanding: There is a distinct lack of mutual understanding of session key in the Park *et al.*'s protocol. However, the proposed login and authentication phase of the proposed protocol now contains messages 3 and 4, whose main purpose is for the user and remote server to prove to one another that they are holding the same session key. Moreover, it can be seen in the proposed protocol that no session key or its hash value is transmitted between the user and remote server when proving the possession of the session key. Instead, the session key is used in the encryption process, which in turn can demonstrate the holding of the session key.

5) Rainbow table attack: The transmission of the hash value of the session key or $h(SK)$ to prove the knowledge of the session key in the authentication phase of the Park *et al.*'s scheme could be the cause of a rainbow table attack. In this stage, an adversary could intercept $h(SK)$ and potentially find a match in the rainbow table. Thus, the actual session key could potentially be known by the attacker. In the proposed protocol, the issue has been mitigated by changing the proving of the knowledge of the session key. The proposed protocol used an encryption mechanism to accomplish the aim. In other words, when proving that the user and remote server hold the same session key, each entity would encrypt message components instead of hashing the key. This way, the recipient of the encrypted message would only have to decrypt it to see whether or not the other entity actual held the same session key.

Table III compares security functionality and features of our proposed protocol with the Park *et al.*'s protocol as well as the Cao-Ge protocol on which Park *et al.* claimed to have improved. The symbol ∘ denotes that the security feature exists in the protocol. The symbol × denotes that the security feature does not exist in the scheme.

It can be seen from Table III above that when comparing security functionality and features of the proposed multi-factor biometric-based remote authentication protocol with the Park *et al.*'s protocol and the Cao-Ge protocol, there are many advancements obtained from our protocol. First of all, both the Park *et al.*'s protocol and our proposed scheme provide mutual authentication and session key agreement between the user C_i and the server S_i during the login and authentication phase. However, the Cao-Ge protocol does not provide any of these features. Although a session key is established in the Park *et al.*'s protocol, there is still a lack of mutual understanding of the key. In other words, although a key is established between the user and the remote server, the server holds no confirmation that the user is in possession of the same key. However, the proposed protocol has fixed the issue. Both entities can now be sure of the possession of the session key of the other entity.

TABLE III
COMPARISON OF SECURITY FEATURES

Features	Cao-Ge	Park *et al.*	Proposed Protocol
Mutual Authentication	×	∘	∘
Session Key Agreement	×	∘	∘
Key Mutual Understanding	×	×	∘
Message Falsification Resistance	×	×	∘
Replay Attack Resistance	×	×	∘
Man-in-the-Middle Attack Resistance	×	×	∘
Rainbow Table Attack Resistance	×	×	∘

Next, message falsification, replay attack and main-in-the-middle attack, which are the vulnerabilities of both the Cao-Ge and Park *et al.*'s protocols, have been addressed by the proposed scheme with the application of a cryptographic hash function and challenge-response mechanism. Finally, there is the possibility of a rainbow table attack in the Park *et al.*'s protocol due to the way that the server proves the knowledge of the session key to the user, which is by transmitting the hash value of the session key. This problem has been mitigated by the proposed protocol. Encryption of message components with the session key is used to prove the possession of the key by both the user and server. This way there is no need to transmit the hash value of the key. Hence, a rainbow table attack is mitigated by the proposed protocol.

C. Performance analysis

This section compares the computational cost between the Cao-Ge protocol, Park *et al.*'s protocol and the proposed multi-factor biometric-based remote authentication protocol. The following notations were used by Park *et al.*, so it has been decided that the same notations will be used for consistency. Firstly, T_h denotes the computation time for a cryptographic hash function $h()$. Secondly, T_H denotes the computation time for a Bio-Hashing function $H()$. The XOR ($\oplus$) operation is not considered because, according to Park *et al.*, comparing with T_h it can be ignored. The comparison of computational costs of the two scheme are shown in Table IV. Note that the label RC represents the registration centre.

It can be seen that the computational cost of the proposed protocol is very similar to that of the Park *et al.*'s protocol, but slightly different from that of the Cao-Ge protocol. First of all, it takes the user one Bio-Hash operation to complete the registration process in the Park *et al.*'s scheme,

TABLE IV
COMPARISON OF COMPUTATIONAL COST

Process	Cao and Ge			Park *et al.*			Proposed Protocol		
	User	RC	Server	User	RC	Server	User	RC	Server
Registration	0	$7T_h$	0	$1T_H$	$10T_h + 3T_H$	0	$1T_h + 1T_H$	$11T_h + 3T_H$	0
Login and Authentication	$8T_h$	0	$4T_h$	$24T_h + 8T_H$	0	$8T_h + 2T_H$	$26T_h + 8T_H$	0	$5T_h$
Total	$8T_h$	$7T_h$	$4T_h$	$24T_h + 9T_H$	$10T_h + 3T_H$	$8T_h + 2T_H$	$27T_h + 9T_H$	$11T_h + 3T_H$	$5T_h$

while it takes one secure hash operation and one Bio-Hash operation in our proposed protocol. The user is not required to compute any hash operation in the Cao-Ge protocol. In the same process, it takes the registration centre ten secure hash operations and three Bio-Hash operations in the Park *et al.*'s scheme, while it takes eleven secure hash operations and three Bio-Hash operations in the proposed protocol. Meanwhile, it takes the registration centre seven secure hash operations in the Cao-Ge protocol. This means that on the whole, the proposed protocol requires more secure hash operations in the registration process than the other two schemes. This is because the extra hash operations from the added cryptographic hash function are required for the checking of message integrity in the messages of the registration phase.

The second process to be considered is the login and authentication phase. It takes the user twenty-four secure hash operations and eight Bio-Hash operations to complete the process in the Park *et al.*'s protocol, while it takes twenty-six secure hash operations and eight Bio-Hash operations in our proposed scheme. At the same time, the vulnerable Cao-Ge protocol requires the user to compute eight secure hash operations during the login and authentication process. The proposed scheme contains more secure hash operations due to the extra messages used to ensure the mutual understanding of the session key.

Moreover, it takes the remote server four secure hash operations in the Cao and Ge's scheme, eight secure hash operations and two Bio-Hash operations in the Park *et al.*'s protocol, while it *only* takes five secure hash operations in our protocol. Comparing the proposed protocol with the insecure Cao-Ge protocol, the proposed scheme requires one more secure hash operation. When comparing the Park *et al.*'s protocol with the proposed protocol, it appears that the proposed scheme requires less computational cost to accomplish the aims of mutual authentication and key agreement.

VI. CONCLUSION

The Internet has provided many services that have become more an more popular. Unfortunately, it also comes with many security problems, including the issue of access control. One way to reduce the risk is to ensure that only authorised users are allowed to access the service with a secure channel being established.

In this paper, it has been demonstrated that the existing multi-factor biometric-based authentication protocols, specifically the Park *et al.*'s scheme, contain several vulnerabilities. They include the lack of message integrity checking mechanism, the possibility of replay attacks, the lack of a challenge-and-response mechanism, the lack of mutual understanding of the session key and the possibility of a rainbow attack.

An improved protocol believed to be able to overcome the mentioned weaknesses has, therefore, been proposed. In the proposed multi-factor biometric-based authentiation protocol, two messages are needed to carry out the registration phase, while four messages are required to complete the login and authentication phase. The proposed scheme has also been proved and analysed by the logic of GNY.

Moreover, In order to illustrate the improvements accomplished by the proposed protocol, the comparison was done on security functionality and features of the Cao-Ge protocol and the Park *et al.*'s protocol. It has been shown that the proposed scheme can overcome the weaknesses of the Cao-Ge protocol and the Park *et al.*'s protocol. These include the use of a cryptographic hash function for message falsification detection, the use of encryption to mitigate rainbow table attack and the application of challenge-and-response mechanism and nonce values to prevent replay attack, man-in-the-middle attack as well as to ensure the mutual key understanding.

In addition, the computational cost of the proposed multi-factor biometric-based authentication protocol was compared with existing schemes, specifically the Cao-Ge protocol and the Park *et al.*'s protocol. The comparison with the Cao-Ge protocol shows that the proposed protocol requires higher computational cost, but it comes with better security. When comparing the proposed scheme with the Park *et al.*'s protocol, it seems that the proposed scheme requires a very similar computational cost in the registration phase. In the login and authentication phase, on the other hand, it appears that the proposed protocol requires less computational cost.

REFERENCES

[1] C. Lin, H. Sun, and T. Hwang, "Attacks and solutions on strong-password authentication," *IEICE Transactions on Communications*, vol. 84, no. 9, pp. 2622–2627, Sep. 2001.

[2] C. Li and M. Hwang, "An efficient biometrics-based remote user authentication scheme using smart cards," *Journal of Network and Computer Applications*, vol. 33, no. 1, pp. 1–5, Jan. 2010.

[3] A. Das, "Analysis and improvement on an efficient biometric-based remote user authentication scheme using smart cards," *IET Information Security*, vol. 5, no. 3, pp. 145–151, Sep. 2011.

[4] Y. An, "Security analysis and enhancements of an effective biometric-based remote user authentication scheme using smart cards," *Journal of Biomedicine and Biotechnology*, vol. 2012, pp. 1–6, 2012.

[5] L. Cao and W. Ge, "Analysis and improvement of a multifactor biometric authentication scheme," *Security and Communication Networks*, vol. 8, no. 4, pp. 617–625, May 2015.

[6] Y. Park, K. Park, K. Lee, H. Song, and Y. Park, "Security analysis and enhancements of an improved multi-factor biometric authentication scheme," *International Journal of Distributed Sensor Networks*, vol. 13, no. 8, pp. 1–12, Aug. 2017.

[7] L. O'Gorman, "Comparing passwords, tokens, and biometrics for user authentication," *Proceedings of IEEE*, vol. 19, no. 12, pp. 2021–2040, Dec. 2003.

[8] A. Bhargav-Spantzel, A. Squicciarini, S. Modi, M. Young, E. Bertino, and S. Elliott, "Privacy preserving multi-factor authentication with biometrics," *Journal of Computer Security*, vol. 15, no. 5, pp. 529–560, 2007.

[9] S. Boonkrong, "Internet banking login with multi-factor authentication," *KSII Transactions on Internet & Information Systems*, vol. 11, no. 1, pp. 511–535, Jan. 2017.

[10] W. Simpson and K. E. Foltz, "Secure identity for enterprises," *IAENG International Journal of Computer Science*, vol. 45, no. 1, pp. 142–152, 2018.

[11] L. L. Gong, R. Needham, and R. Yahalom, "Reasoning about belief in cryptographic protocols," in *Proceedings 1990 IEEE Symposium on Research in Security and Privacy*. IEEE Computer Society Press, 1990, pp. 234–248.

[12] M. Burrows, M. Abadi, and R. Needham, "A logic of authentication," *ACM Transactions on Computer Systems*, vol. 8, pp. 18–36, Feb. 1990.

[13] S. Boonkrong and C. Somboonpattanakit, "Dynamic salt generation and placement for secure password storing," *IAENG International Journal of Computer Science*, vol. 43, no. 1, pp. 27–36, 2016.

[14] X. Wang and H. Yu, "How to break md5 and other hash functions," in *Annual international conference on the theory and applications of cryptographic techniques*. Springer, 2005, pp. 19–35.

Sirapat Boonkrong obtained a Bachelor of Science Degree in Computer Science in 2002 and a Doctor of Philosophy in Computer Science in 2006 from the Univesity of Bath, UK. His Ph.D. thesis was titled "Authenticaion, Pre-Handoff and Handoff in Pure MANET", for which he received the Dissertation Award from the National Research Council of Thailand (NRCT).

He began his career by being a researcher at the National Electronics and Computer Technology Centre, Thailand. Between August 2009 and August 2017, he was a full-time lecturer at the Faculty of Information Technology, King Mongkuts University of Technology North Bangkok (KMUTNB) mainly teaching in the field of information and computer network security to both Master and Ph.D. students. His final management position at KMUTNB was an Assistant to the President for Research Affairs and Information Technology.

Currently, he is a full-time lecturer at the School of Information Technology at Suranaree University of Technology (SUT) and is still teaching and researching in the area of information security. He is also dolsing a position of Deputy Director of the Centre for International Affairs at the university.